SPATIAL-TEMPORAL ANALYSIS OF OPENSTREETMAP DATA AFTER NATURAL DISASTERS: A CASE STUDY OF HAITI UNDER HURRICANE MATTHEW

Jin Xu , Longxi Li, Qi Zhou*

Faculty of Information Engineering, China University of Geosciences, Lumo Road, Wuhan, P.R.China – xujinxgem@outlook.com

Commission IV, WG IV/3

KEY WORDS: OpenStreetMap, Spatial-temporal Analysis, VGI, Crisis Mapping, Hurricane Matthew

ABSTRACT:

Volunteered geographic information (VGI) has been widely adopted as an alternative for authoritative geographic information in disaster management considering its up-to-date data. OpenStreetMap, in particular, is now aiming at crisis mapping for humanitarian purpose. This paper illustrated that natural disaster played an essential role in updating OpenStreetMap data after Haiti was hit by Hurricane Matthew in October, 2016. Spatial-temporal analysis of updated OSM data was conducted in this paper. Correlation of features was also studied to figure out whether updates of data were coincidence or the results of the hurricane. Spatial pattern matched the damaged areas and temporal changes fitted the time when disaster occurred. High level of correlation values of features were recorded when hurricane occurred, suggesting that updates in data were led by the hurricane.

1. INTRODUCTION

Great damages were made by natural disasters every year due to the inefficiency of data updates. To fill up the blank of delayed authoritative data, Volunteered geographic information (VGI) was adopted as a solution where OpenStreetMap (OSM) played an important role. The online collaborative mapping community had proved to be useful during the 2010 Haiti Earthquake, and it is now aiming at conducting crisis mapping ever since (Yuma Imi et al.2012).
A better understanding was required to select information more efficiently from OSM database after natural disasters occurred. Spatial-temporal pattern was needed to better understand how OSM data was affected by natural disasters as they occurred so suddenly that significant updates might be observed. In previous studies, roads and buildings were considered to be the major features updated by contributors, thus roads and buildings were analyzed for answering the following questions:

1. How did the OSM road lengths and OSM building number increase before and after disasters happened?
2. And how did the length of updated road data and the number of updated building data in OSM vary in both spatial and temporal dimensions?
3. Was there a correlation between the length of updated roads and the number of updated building data in OSM? If there was, how did these correlations vary before and after disasters?

Hence, we chose Haiti as an example, as it was struck by Hurricane Matthew in October, 2016. We downloaded OSM data of the Republic of Haiti during which time the nation was tremendously damaged by Hurricane Matthew. The nation is subject to hurricanes under the tropical climate and was struck by the Category 4 hurricane during October 4th to October 5th in 2016 while it is still recovering from the 2010 earthquake. Another reason why Haiti was chosen as the study area is that its poor infrastructure conditions and poverty make it even more concerning when responding to natural disasters in a short time to prevent further damages.

2. RELATED WORKS

Disasters normally are singular large scale, high impact events (Susan L. Cutter. 2003). Thus damages were made, leading to great losses in lives and properties. In the era of big data, crowdsourced data and VGI are now the huge source of data (Alexander Fekete et al. 2015) that was considered to provide an alternative to traditional authoritative information (Micheal F. Goodchild, et al. 2010). As VGI became a promising source in time-critical situations (Micheal F. Goodchild, et al. 2010) and its initiative enhances the disaster response and improves the results in post-disaster recovery (Luiz A. Manfré et al.2012), many studies in VGI and disasters were conducted.
In some previous studies, data containing geographic information obtained from social network such as Tweeter were analyzed. Tweeting were vastly concentrated in areas hit hardest by the Hurricane Sandy (Taylor Shelton et al.2014) and population of re-tweeting activity in geographically vulnerable places differs from that of the general (Marina Kogan et al. 2015). Temporal changes of social media data after Typhoon Ruby hit Philippines were also analyzed, whose results showed that more than 76% of all contributions were streamed during the event. (Mohammad Ebrahim Poorazizi et al. 2015). Although these studies indicated that VGI contributors are much more active in response to disasters, geographic information whose contributors are local users is not as developed in some areas.
Luckily, society has now entered a new era where geographic information will not only be used but also be created by observers (Micheal F. Goodchild, et al. 2010). Sources of geographic information are not only local users but also foreign volunteers. In the case of OSM, rapid increase of registered users brought in updates in data (Muki Haklay et al. 2014), whose contributors may derive a certain personal satisfaction in the opportunity to create and publish VGI (Micheal F. Goodchild. 2007). When disasters occurred, contributors might be more motivated due to the great personal satisfaction and attractions caught by the tragedies.
As OSM being a large community of active contributors for collecting geographic information by volunteers around the

world (Pacle Neis et al. 2014), many studies were conducted in recent years, but few case studies were about spatial-temporal changes caused by natural disasters.

In one study, a common law was discovered in both spatial and temporal changes of OSM data (Tzu-Yuan Chung et al. 2015), but how does the vogue concept present was still unclear. In another study, geographical events were taken into account as a factor for building a spatial-temporal VGI model (Yijiang Zhao et al. 2016), but it focuses more on how to build these models without much study about spatial-temporal changes of OSM data. The number of user-generated placemarks indexed by Google Map was studied in another research for spatial changed (Matthew A. Zook et al. 2010) and the results showed that placemarks were updated more in affected areas after natural disasters. However, does OSM data have the same feature?

Similar studies were also conducted for temporal changes of many features after the great east Japan earthquake (Yuma Imi et al.2012), whose result suggested a significant update in roads and buildings. Spatial-temporal analysis was conducted in a study to prove that OSM data will response to natural disasters (Thiago Henrique Poiani et al. 2016). Number of new and active OSM data contributors was recorded and spatial analysis was conducted during crisis.

Although previous studies had given some ideas of how OSM data changes, further studies were still needed to understand how did the OSM road lengths and OSM building number updated before and after disasters happened in both temporal and spatial dimension, whether correlation existed, and how correlation changed in time dimension if it existed was still unknown.

3. RESEARCH DESIGN

We downloaded 18 days of OSM data from OpenStreetMap Data Extracts (http://download.geofabrik.de/index.html), where OSM data will be uploaded daily and can be freely downloaded. In order to figure out the difference, we downloaded data from October 1st to October 18th, covering the time before and after Hurricane Mathew struck Haiti.

The raw data consisted of 18 features including roads, buildings, landuse, natural lands, places, pofw, pois, railways, traffic, transport, water, and waterways. Features of buildings and roads were analyzed in this study as major features that attracted great attention from contributors.

We also downloaded administration maps from Global Administrative Areas (http://www.gadm.org/dowload). The administration used in this study was commune, which is commonly used in spatial study in Haiti. This administration with 134 subjects in its attribute table was further divided from the administration of districts, ranking forth in the administration class while the first was the whole country, the second was departments, and the third was districts. The administration was shown in Figure 1.

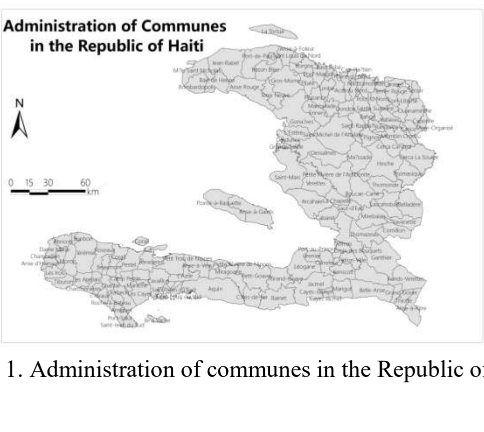


Figure 1. Administration of communes in the Republic of Haiti

3.1 Temporal Analysis

First we analyzed the quantity of OSM data updated after the hurricane. In order to know how long roads were updated and how many buildings were updated, we used the tool "Erase" from ArcToolBox to erase data of one day earlier from current data. For example, data of October 1st was erased from data of October 2nd for updated data that volunteers uploaded on October 2nd.

After that, we analyzed how many changes had been made daily by calculating the length of roads and number of buildings updated in Haiti every day. Daily updates were recorded to draw a figure of the temporal changes at a rate of one day for understanding whether temporal changes of data were related to natural disasters and how data updates were influences by disasters if related.

3.2 Spatial-temporal Analysis

Meanwhile, we also analyzed how updates of OSM data scattered by locating hotspots where data was mostly updated every day.

Communes with the most updates were considered to be spatial indicators for areas which might be severely damaged or affected by the hurricane. These areas were classified by statistic results of OSM daily updates. Length of roads and number of buildings updated in every commune every day were recorded. The results will later be used in classifying 3 different types of updates level.

Polygon type as buildings was offered in raw data, it was needed to turn this feature into points for a fair comparison before calculating the updated quantity. New fields were added to attribute table to calculate how long the roads were updated and how many buildings were updated in every communes in Haiti. The result was used for classification.

Recorded length of roads and numbers of buildings were classified into 4 colours representing different levels of updated quantity. Thresholds for classification were based on geometric interval classification methods that can be calculated with the following equation:

$$a_n = a_1 r^{n-1} \tag{1}$$

where a_n = thresholds for classification
a_1 = the first threshold
r = common ratio

The actual values of each terms mentioned above were determined by statistic histograms of daily updated OSM data who indicated the interval of data quantity.

Thresholds were applied in cartography and after cartography, names of hotspots were recorded. Media reports were also collected to figure out whether spatial changes have anything to do with Hurricane Mathew.

3.3 Correlation Analysis

We also analyzed the correlation between the two main features in order to figure out whether there is any correlation between the spatial-temporal changes the OSM updates of length of roads and number of buildings. Pearson correlation coefficient was the criterion to understand the correlation.

Pearson correlation coefficient values of how the updated length of roads and number of buildings scattered in spatial dimension every day were calculated in a way of figuring out whether correlation existed between these two features. The results of

https://doi.org/10.5194/isprs-archives-XLII-2-W7-1401-2017

daily correlation coefficient values were recorded in order to understand how it changed in temporal dimension and if natural disaster played a part in these changes.

4. RESULTS

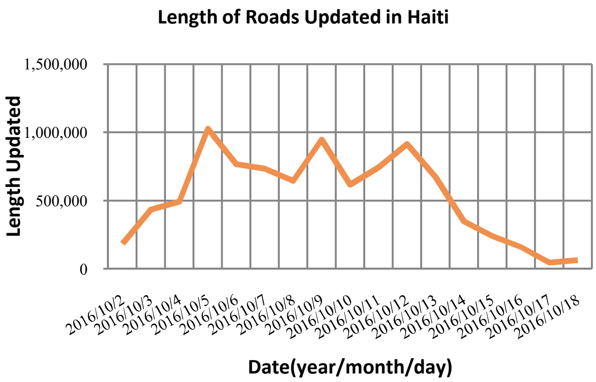
4.1 Results of Temporal Analysis

17 days of OSM data of roads and buildings in Haiti was recorded by their total updates in temporal dimension, whose result was drawn as below (Figure 2).

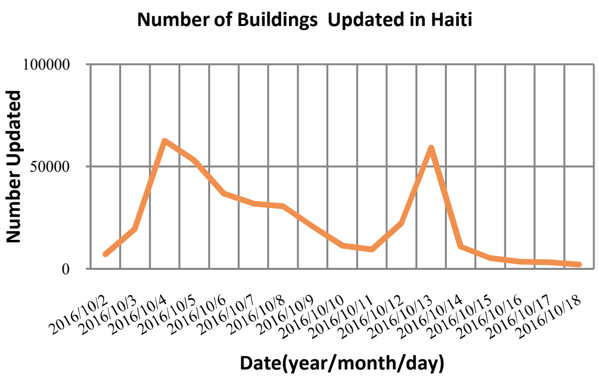
According to the result, the first significant updates had been made by contributors at about October 4th, and fewer updates were uploaded before October 4th and after October 13th. During this time, roads updates were generally at a high level while buildings showed two significant peaks. Massive updates were recorded in the following 4 days after the hurricane and gradually deduced to a low level afterwards.

Updates of roads were generally at a high level during the time from October 5th to October 12th, with a certain fluctuation. Updates of buildings presented two obvious peaks at October 4th and October 13th. Updates of buildings remained at a high level for about four days and these updates slowly returned to a low level. Before the first significant update, both features indicated slow growths in data. However, after October 13th, the updates of both roads and buildings slowly returned to a low level.

The first Peak of buildings appearing at around October 4th matched the time when Hurricane Matthew hit Haiti. Natural disasters caught attention of many contributors on OSM, resulting in the first significant update of OSM data. After the first peak, contributors were still interested in affected areas and kept updating, leading to the lasting high level of updates.



(a)



(b)

Figure 2. Temporal changes of OSM data updated as roads (a) and buildings (b).

The update at about October 13th was long after the hurricane struck Haiti, but it indicated that data contributors were still interested in the affected area. The peak of buildings suggested that it might be another event leading to another attention in Haiti. As October 13th 2016 being the International Day for Natural Disaster Reduction, aiming at reducing the harm of natural disasters, attention of contributors was caught again because Haiti was struck by hurricane recently.

Roads showed a feature of high level of updates in general during October 5th to October 12th, suggesting that OSM data contributors were more interested in roads than buildings. Another evidence is the quantity of data updated before Hurricane Matthew hit Haiti, during which time roads showed a growth in updates while buildings grew more slowly. This phenomenon is obvious on October 3rd. The fact that roads were much easier to draw on map was probably the main reason.

A sign of updates before Hurricane Matthew hit Haiti was also noticed. As hurricane can be predicted, it is possible for contributors to upload OSM data before natural disaster happened. Reports of official warnings were published few days before Hurricane Matthew hit the nation, supporting the theory that the sign of updates were related to the warning.

4.2 Results of Spatial-temporal Analysis

We defined 2km, 10km and 50km as the thresholds for classifying the length of roads while 200, 1000 and 5000 were the thresholds for classifying number of buildings. The first level for classification was spatial indicator pointing out the most updated areas, where damages were assumed to be severe. The second level indicated that these areas attracted much attention from OSM data contributors, which were also assumed to be damaged but in a less severe condition. The last two levels suggested that these areas also attracted interests of OSM data contributors, which were assumed to be affected by the hurricane but were less damaged than those areas classified into the first and the second level.

Areas with significant updates of both roads and buildings were classified, made into maps and shown in Figure 4 and 3.

These maps together showed a certain pattern of routes crossing through Haiti from the south to the northwest of the nation. Both features appeared in clustering states in the southwest before the hurricane struck Haiti, and these states changed after the hurricane occurred.

Roads were mainly updated in a few communes in the southwest of Haiti before October 3th when this feature was updated in a small amount. At October 4th, hotspots appeared mainly in the southwest of the country. Some areas in other parts of the nation were also recorded for a certain quantity of updates. Rapid updates of roads during October 5th to October 12th were mainly scattered in the southern part of the country and updates were recorded in a larger range than before. Many communes were classified as hotspots during this time and they were mainly in the south. After October 12th, data was mostly updated in the northwest with a small amount in the south. Moving of high-lighted areas indicated the shifting of attention of OSM data contributors. Communes with massive updates were more separated after October 12th, the number of which had a slight reduction every day. After October 15th, OSM data still had a trend of updating but the updates were generally at a lower level and hotspots for updated length of OSM roads were much fewer than before.

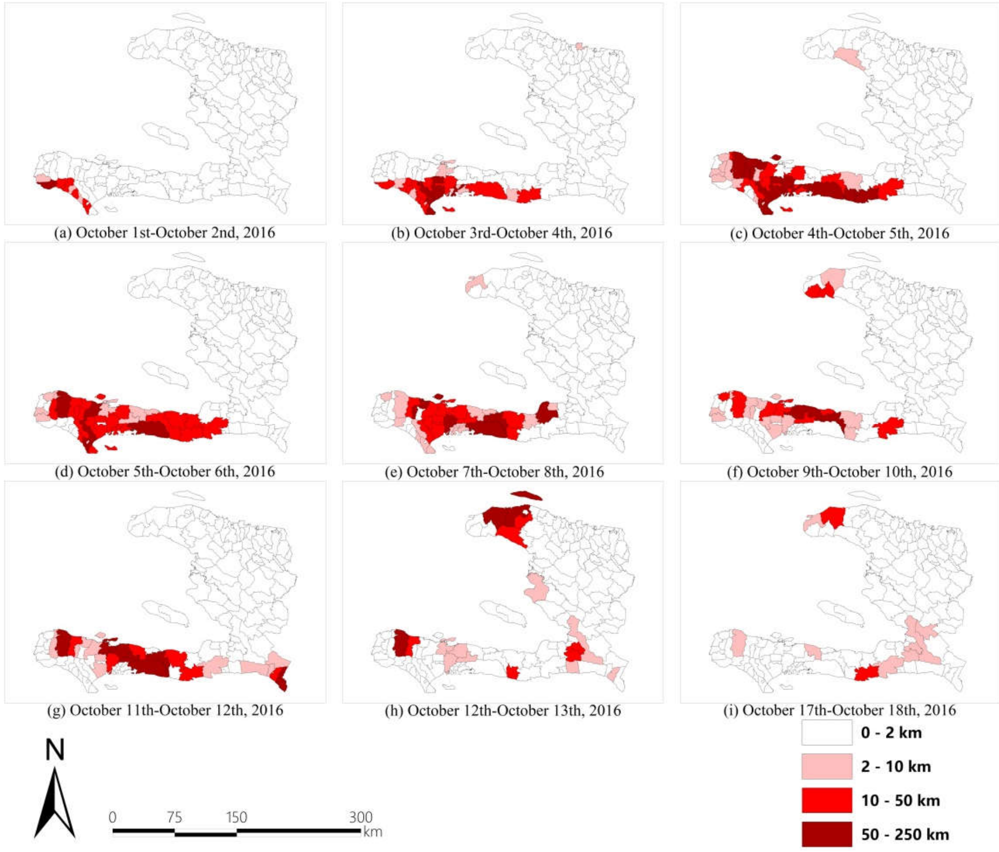


Figure 3. Length of roads updated in every commune in Haiti.

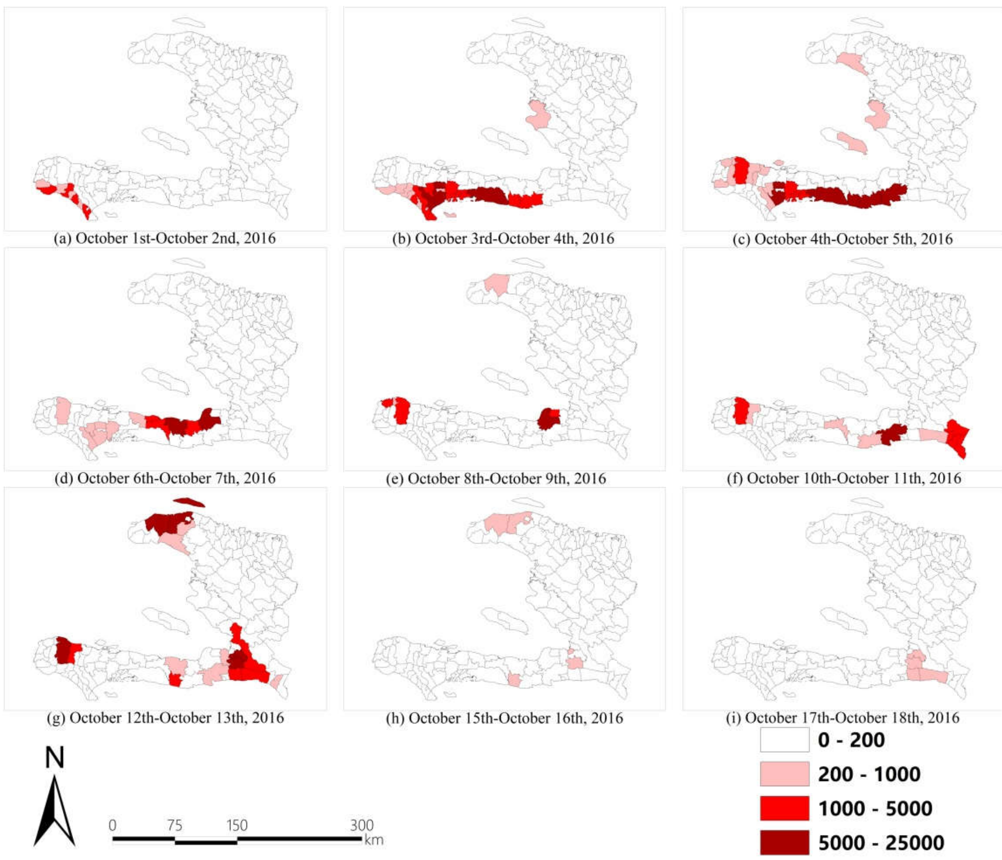


Figure 4. Number of buildings updated in every commune in Haiti

Buildings showed a similar pattern for updates which located mostly in the south and in the northwest. Before October 4th when buildings were recorded for a significant update in total data, it was updated in few communes in the southwest. During October 4th to October 13th, more hotspots appeared and they were mainly in the south with a few in the north. Most communes were updated with small numbers of buildings, while the minority were updated in huge amounts of OSM data, indicating that some areas attracted more attention from data contributors. After October 13th, OSM data still showed a trend of updating, however the quantity of updates in buildings were much smaller and areas with updates separated in the northwest and the south.

Both features also showed an expansion in spatial ranges of hotspots after the hurricane stroke Haiti. More communes were updated for OSM data after the first significant updates in temporal dimension whereas updated areas were clustered in the southwest before. Hotspots were mainly in the south with a few in the northwest, and some communes in the southwest were frequently marked as hotspots.

Meanwhile, media reports were gathered locating the damaged areas to confirm whether OSM data was updated in affected areas. According to the media, severe damages were made in the south and the northwest due to the hurricane (http://news.xinhuanet.com/world/2016-10/10/c_129315554). The north of the country was greatly damaged (http://www.bbc.com/news/world-latin-america-37596222).

Hotspots of OSM updates matched the reports in general, indicating a relation between the most affected areas and OSM data updates. Areas with tremendous damages were also shown in maps as hotspots with huge amounts of updated data, indicating that the attention of OSM contributors was caught by the hurricane.

4.3 Results of Correlation Analysis

OSM data of roads and buildings were recorded by time and communes, which were analyzed for correlation. The result of Pearson correlation coefficient value was shown in Figure 5.

According to the scatter diagram, correlation coefficient values between roads and buildings were above 0.6 during October 1st to October 5th as well as October 12th to October 15th while correlation values were relatively low during October 6th to October 11th. The Pearson correlation coefficient value in October 9th was 0.136, and it was noted as uncorrelated.

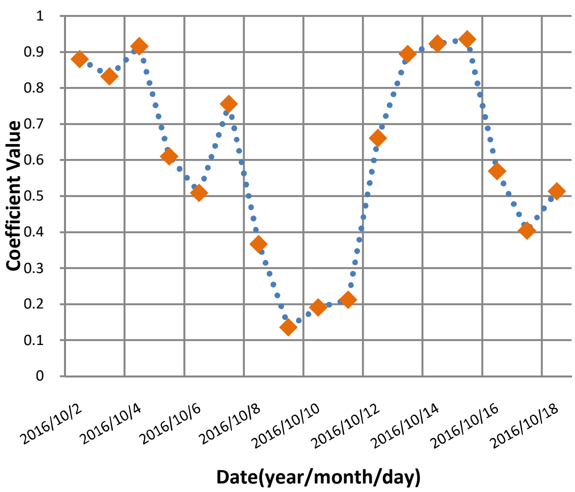


Figure 5. Pearson correlation coefficient value between roads and buildings

Generally, correlation between roads and buildings existed and it was at a moderately high level when the hurricane hit Haiti while it gradually dropped to a lower level within a week.

The result indicated that changes of OSM data were the result of Hurricane Matthew as roads and buildings became linear dependent at the time when the hurricane struck the nation while the correlation coefficient values went down to a low level when the hurricane was gone. The moderately high level of Pearson correlation coefficient values suggested a similarity in spatial-temporal changes of OSM data, indicating these two features were updated together due to the attention from data contributors.

Pearson correlation coefficient values were also at a high level before the hurricane arrived, which can be easily understood as both features were updated in the southwest part of the nation, crowing in a small area with little updated data. During October 6th to October 11th, Pearson correlation coefficient values were mostly under 0.6, which can be explained as small amounts of updated data was scattered in a larger area. The sudden update of correlation value around October 13th was mainly caused by the International Day for Natural Disaster Reduction as this special day drew the attention again after Haiti being damaged by natural disaster few days ago.

5. CONLUSIONS AND DISCUSSIONS

The result of this study suggested a sudden interest of data contributors led by the hurricane. Their behaviour of updating OSM data of roads and buildings was under the influence of Hurricane Matthew, resulting in the significant updates shortly after disaster occurred together with the significant updates in affected areas. Roads and buildings showed correlation between each other especially at the time when Hurricane Matthew hit Haiti and International Day for Natural Disaster Reduction.

OSM data of roads and buildings were updated rapidly after the hurricane, suggesting sudden attention of contributors which lasted for about 4 days after Hurricane Matthew hit Haiti. Roads attracted more attention from OSM contributors than buildings and was updated in a larger quantity. However, only 18 days of OSM data were downloaded in this study, further studies were still required to obtain more historical data.

OSM data of roads and buildings were mainly updated in the damaged areas. Affected areas were also updated but in a smaller quantity. Although maps were capable of being a spatial indicator from the global scale, accuracy of pointing out the exact affected communes remaining low. Moreover, the lacking of reports made it harder to further confirm whether OSM data updates matched the damages in a smaller scale.

Correlation of spatial changes existed and was at a moderately high level when the hurricane struck Haiti and about the time of International Day for Natural Disaster Reduction, while the values were low during ordinary days. The result also indicated that updates in OSM data of roads and buildings were the result of the hurricane. However this study only studied liner dependence whereas whether non-linear dependence existed would still needed to be further studied.

This study was conducted in underdeveloped areas. Further study will be conducted in developed areas with more complete OSM data.

ACKNOWLEDGEMENTS

The project was supported by National Natural Science Foundation of China (No. 41301523) and Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (No. G1323541711).

REFERENCES

Cutter S L., 2003. GI Science, Disasters, and Emergency Management. Transactions in Gis, 7(4),pp.439-446.

Chung T Y, Chuang K T, Hsu C M, et al., 2015. Spatiotemporal Crowdsourcing Behavior: Analysis on OpenStreetMap. In: Conference on Technologies and Applications of Artificial Intelligence, IEEE, pp.373-380.

Fekete A, Tzavella K, Armas I, et al., 2015. Critical Data Source; Tool or Even Infrastructure? Challenges of Geographic Information Systems and Remote Sensing for Disaster Risk Governance. Isprs Journal of Photogrammetry & Remote Sensing,4, pp.1848-1869.

Goodchild M F., 2007, Citizens as Sensors: the World of Volunteered Geography. GeoJournal, 69(4), pp.211-221.

Goodchild M F, Glennon J. A., 2010. Crowdsourcing Geographic Information for Disaster Response: A Research Frontier. International Journal of Digital Earth, 3(3), pp.231-241.

Haklay M, Weber P., 2008. OpenStreetMap: User-Generated Street Maps. Pervasive Computing IEEE, 7(4),pp.12-18.

Imi Y, Hayakawa T, Ito T, et al., 2012. Analyzing the Effect of Open Street Map During Crises: The Great East Japan Earthquake. In: Congress on Evolutionary Computation, IEEE, pp.126-130.

Kogan M, Palen L, Anderson K M., 2015. Think Local, Retweet Global: Retweeting by the Geographically-Vulnerable during Hurricane Sandy. In: ACM Conference on Computer Supported Cooperative Work & Social Computing. ACM, pp.981-993.

Manfré L A, Hirata E, Silva J B, et al.,2012. An Analysis of Geospatial Technologies for Risk and Natural Disaster Management. ISPRS International Journal of Geo-Information, 1(2), pp.166-185.

Neis P, Zielstra D., 2014. Recent Developments and Future Trends in Volunteered Geographic Information Research: The Case of OpenStreetMap. Future Internet, 6(1), pp.76-106.

Poiani T H, Rocha R D S, Degrossi L C, et al.,2015. Potential of Collaborative Mapping for Disaster Relief: A Case Study of OpenStreetMap in the Nepal Earthquake 2015. In: Hawaiian International Conference on System Sciences - Hicss, pp.188-197.

Poorazizi M, Hunter A, Steiniger S.,2015. A Volunteered Geographic Information Framework to Enable Bottom-Up Disaster Management Platforms. ISPRS International Journal of Geo-Information, 4(3), pp.1389-1422.

Shelton T, Poorthuis A, Graham M, et al.,2014. Mapping the Data Shadows of Hurricane Sandy: Uncovering the Sociospatial Dimensions of 'Big Data'. Geoforum, 52(52), pp.167–179.

Zhao Y, Zhou X, Li G, et al., 2016. A Spatio-Temporal VGI Model Considering Trust-Related Information . ISPRS International Journal of Geo-Information, 5(2):10.

Zook M, Graham M, Shelton T, et al.,2010. Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake. Social Science Electronic Publishing, 2(2), pp.7–33.